A Scalable Framework for Post Fire Debris Flow Hazard Assessment Using Satellite Precipitation Data

Elijah Orland[a,b], Dalia Kirschbaum[b], Thomas Stanley[a,b]

[a] University of Maryland Baltimore County, GESTAR II, Baltimore, Maryland, United States of America

[b] Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland, United States of America

Corresponding author:

Elijah Orland

(elijah.orland@nasa.gov)

NASA Goddard Space Flight Center, Hydrological Sciences Lab, 8800 Greenbelt Rd, Greenbelt, MD, 20771, USA

Keywords: **Wildfire, Debris Flows, Mass Wasting, Remote Sensing, Machine Learning, IMERG**

Abstract

Wildfire is a global phenomenon that has dramatic effects on erosion and flood potential. On steep slopes, burned areas are more likely to experience significant overland flow during heavy rainfall leading to post fire debris flows (PFDFs). Previous work establishes methods for PFDF hazard assessment, often relying on regional-scale parameterizations with in-situ rainfall measurements to categorize hazard as a function of meteorological and surface properties. We present a globally scalable approach to extend the benefit these models provide to new areas. Our new model relies on publicly available satellite-based inputs with a global extent to provide first order hazard assessments of recently burned areas. Our results show it is possible to identify the conditions relevant for PFDF-initiation processes across a variety of physiographic settings. Improvements to satellite-borne rainfall intensity data and increased availability of PFDF occurrence data worldwide are expected to enhance model skill and applicability further.

Plain Language Summary

Wildfires are destructive hazards whose subsequent effects can be long lasting. This includes the increased risk of flooding and post fire debris flows (PFDFs). We present a framework that uses freely available satellite-based data to inform where PFDF activity may be elevated. We use a standardized set of resources available globally such that it is easier to make direct comparisons between regions. This work lays the groundwork to more thoroughly investigate why PFDFs are more common in some regions over others, as well as demonstrate the use of satellite data for timely assessments of the hazards that follow wildfire.

1. Introduction

Driven by the effects of wildfire, burned steeplands are vulnerable to the cascading hazard of runoff-induced post fire debris flows (PFDFs). These hazards are a consequence of the reduced infiltration capacity of burned soils (Debano, 2000; Letey, 2001) and an increase in available sediment for transport (Florsheim et al., 1991, 2016; Gabet, 2003; Lamb et al., 2011, 2013). The mixing of runoff from heavy rainfall and loose debris can create flows capable of destroying infrastructure and taking the lives of those in the path downslope. As wildfires are already occurring at higher frequencies and severities (Abatzoglou & Williams, 2016; Cannon & DeGraff, 2009; Dennison et al., 2014; Keeley, 2009; Miller & Safford, 2012; Mueller et al., 2020; Westerling et al., 2006), there is a continued need to understand how differences in topography, burn severity, and soil properties may affect intensity thresholds for PFDF initiation.

Previous work details several methods for PFDF hazard assessment at the regional or local scale. These approaches include physically-based rainfall intensity-duration threshold delineation via dimensionless discharge and Shields stress criterion (Tang et al., 2019), and data-driven empirical approaches that assign a probability of PFDF initiation and/or estimate of sediment delivery from field-based or remote sensing-derived preconditions (e.g., Gartner et al., 2014; Nikolopoulos et al., 2018; Nyman et al., 2015; Staley et al., 2016). The United States Geologic Survey (USGS) implements a version of the logistic regression model from Staley et al. (2016) as an operational tool to derive rainfall intensity thresholds for PFDF initiation within the Contiguous United States (CONUS). This model is trained on a database compiled by the USGS that documents PFDF occurrence (or the lack thereof) and the associated storm characteristics concurrent with that observation since 2000. Empirical models, such as the one utilized by the USGS, remain a popular approach for PFDF hazard assessment as numerical or physically based models traditionally rely on site-specific parameterizations. These data-driven methods link heterogenous empirical observations across physiographic regions to assign spatially variable probabilities of PFDF occurrence. From an operational standpoint, empirical approaches provide a computationally efficient and transferrable method appropriate for rapid hazard assessment, even if the entirety of PFDF initiation processes cannot be fully derived from the chosen model inputs alone.

Since 2016, NASA's Goddard Space Flight Center has run the near-global *LHASA* (Landslide Hazard Assessment for Situational Awareness) model (Kirschbaum & Stanley, 2018; Stanley et al. 2021). This model provides near-real time assessments of landslide hazards worldwide with approximately 3hrs latency. To date, this model has not explicitly incorporated the impacts of wildfires on mass wasting processes into its assessment framework. Following the ever-increasing importance of wildfires and their subsequent effects, this work describes a complementary model for PFDF hazard susceptibility to provide physically interpretable assessments to burned steeplands at the same scale and timeframe.

Further, this work details the first PFDF hazard assessment model solely from publicly available, global scale remote sensing data and demonstrates the utility that it can provide for vulnerable communities and infrastructure. Although PFDF activity is well documented and actively monitored in the United States, PFDFs also occur in other regions (e.g, García-Ruiz et al., 2013; Jin et al., 2022; Nyman et al., 2019). To our knowledge, no work has attempted to reconcile the entire spatial extent of these events to quantify and dynamically monitor PFDF hazard potential in all areas where they occur. We thus propose a framework to provide situational awareness of these events in near real-time.

2. Methods

Our proposed model relies on PFDF occurrence data provided by the USGS. Staley et al. (2016) highlight a comprehensive database documenting the relevant slope, burn severity, and soil properties of several burned watersheds across the Western United States. This database further records the intensity-duration characteristics of select storms that occurred in each area, indicating whether a debris flow initiated from each storm. This is the most extensive database of PFDF activity (n = ~1500) that is publicly available, and it serves as the basis for the current model in use by the USGS. As this database consists of field observations, local rain gauge data, and remote sensing using CONUS-specific datasets, we utilize the information contained therein as a guide for replicating the data collection process using only satellite-derived resources with a worldwide extent. Taken together, these publicly available products help recreate the training database provided by the USGS in a format sufficient for the needs of this study. Notably, we do not include soil properties used in the original USGS model. While global soil property products exist, they are inconsistent in both accuracy and compilation method. Accordingly, we elect only to include the extensively tested datasets provided by NASA and the USGS, as these products must meet a consistent set of standards before publication (https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy/data-levels).

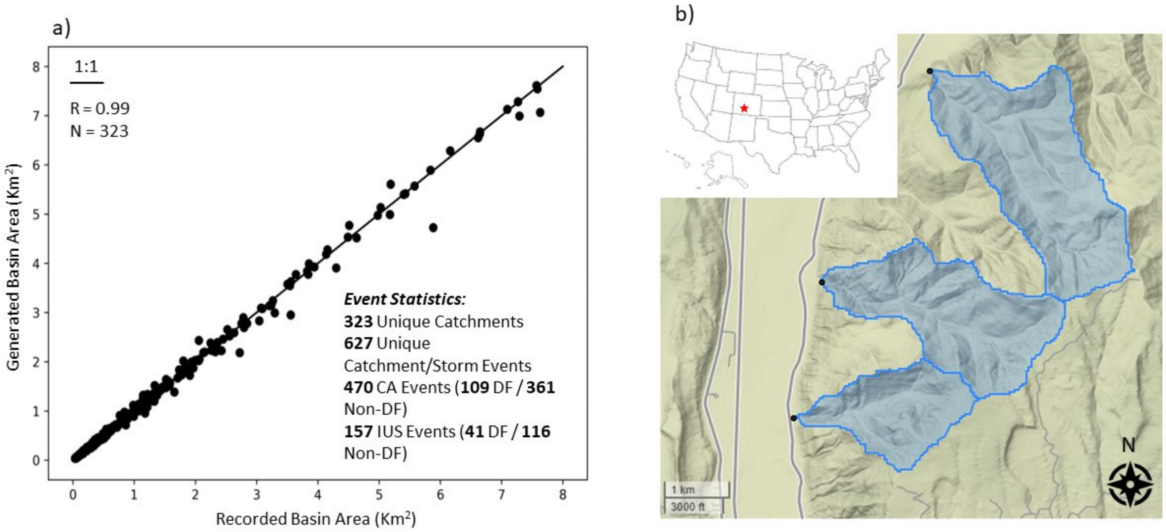

Figure 1. (a) Comparison of catchment areas recorded by Staley et al. (2016) and the catchment areas generated in this work. (b) Example catchments derived in this work (blue) and their corresponding point reference via Staley et al (2016) (black) located near Durango, Colorado. IUS is defined as the Interior United States; CA is defined as Southern California.

An important consideration in the database recreation process is the availability of precise spatial reference information of each documented catchment. The data made publicly available by Staley et al. (2016) link detailed catchment and storm occurrence data to coordinates typically located near basin and/or channel outlets. To relate new observations to the same original assessment areas, it is therefore necessary to delineate catchment boundaries based on the location of each datapoint. We infer catchment boundaries using flow routing software provided by Bartos (2020) and the NASA DEM 30m Digital Elevation Model (NASA JPL, 2020) as a source dataset. The contributing area of each newly generated catchment is compared with the originally recorded value assigned to each point, where both values must be within approximately 20% agreement. This threshold is chosen from manual inspection to ensure that the inferred areas are a reasonable estimation of the original assessment area. Figure 1a provides a quantitative comparison of the resulting catchments areas within this similarity threshold (n = 323). Visual analysis suggests that these catchments are realistic (Figure 1b).

Following delineation, each catchment provides the spatial footprint for a data collection process that replicates the original structure of that provided by the USGS database, and follows a modified approach suitable for the automated, near real-time workflows required by the LHASA model. This includes estimating burn severity based on cloud-free composites derived from Landsat Surface Reflection (SR) products (U.S. Geological Survey) based on the methodology proposed by Parks et al. (2018) using the Google Earth Engine (GEE) platform (Gorelick et al., 2017). Burn severity is quantified by the differenced or "delta" Normalized Burn Ratio (dNBR) (Key & Benson, 2006). Higher dNBR values indicate higher burn severities, while lower or negative values indicate a low severity burn or vegetation regrowth, respectively. For topographic assessment, we use the NASA DEM 30m product as specified earlier.

The most significant departure from the original methodology of Staley et al (2016) is the use of satellite derived rainfall intensity estimate data in contrast to the gauge data in the original database. This also represents one of the major scientific contributions of our work, as we provide the first comparison between satellite and gauge-based intensity thresholds for PFDF initiation in the Western United States. Specifically, we collect 30min max intensity estimate data for each of the storms in the recreated database from the Integrated Multi-Satellite Retrievals for GPM (IMERG) Late Run v06B dataset (Huffman et al., 2019). This 30min peak intensity window is the highest temporal resolution available to us. While 30min data is shown to be a relevant window for PFDF hazard assessment (Liu et al., 2022), we note that PFDF initiation processes are frequently documented to be more effectively delineated by peak intensities at smaller intervals such as 15min or lower (e.g. Cannon et al., 2008; Kean et al., 2011; Staley et al., 2013; Staley et al., 2020). We acknowledge this difference and provide a direct comparison of model performance in the following section.

Following the collection of catchment burn severity, topography, and storm intensity characteristics, we train our probabilistic model using the machine learning algorithm *XGBoost* (Chen & Guestrin, 2016)—a tree-based ensemble model which uses a series of shallow decision trees to iteratively correct residual errors made by the trees before it. This choice contrasts with the logistic regression model detailed in Staley et al. (2016); however, we select the XGBoost algorithm due to its non-linear modeling capabilities and flexibility for setting monotonic constraints that dictate how the probability of PFDF occurrence should correspond to an increase or decrease of a particular variable. These constraints provide physically consistent results such that an increase in rainfall intensity should indeed correspond to an increase in the probability of debris flow initiation. The combination of these features allows for a robust, data driven modeling process to derive complex physical relationships from potentially noisy observations.

To train and validate our model, we rely on a stratified K-Fold cross validation process (He & Ma, 2013) to minimize sampling bias and preserve original class balance. We also recreate the original training/testing scenario documented by Staley et al (2016) for closer comparison between the two models. This approach uses the data collected in California as a training dataset, and model performance is tested on data collected in the Interior United States (IUS). For both training scenarios we report model skill using the Area Under the Curve (AUC) metric, and the Threat Score (TS) which served as the central evaluation metric in Staley et al. (2016), defined as:

$$TS = \frac{True\ Positives}{True\ Positives + False\ Positives + False\ Negatives}.$$

A model with a high TS will demonstrate skill for correctly identifying the conditions most relevant to PFDFs while minimizing classification error.

We finally demonstrate the potential for real time implementation with a case study to evaluate how the model performs on data collected outside of the initial database. This process accomplishes the important tasks of establishing that the model is adequately learning from the database itself, in addition to evaluating model potential in burned steeplands outside of the United States.

3. Results

Figure 2 provides a direct comparison between satellite and gauge derived intensity thresholds for our available training data (n = 627). For IMERG-derived rainfall intensity estimates, the full

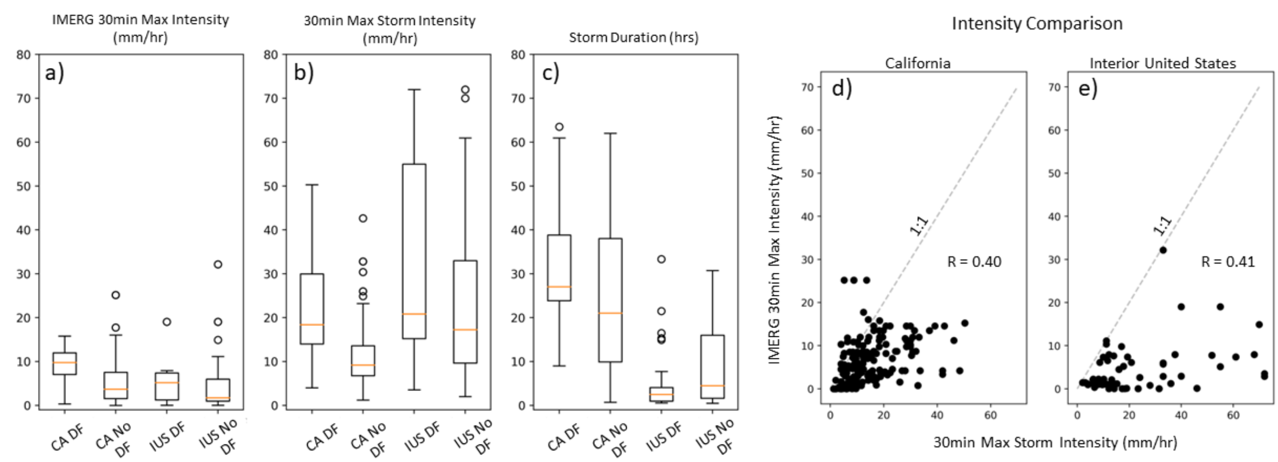

Figure 2. (a) IMERG peak 30min intensity values for DF and Non-DF events separated by region. (b) Gauge-recorded peak 30min intensity values for the same catchments. (c) Storm duration comparison for the intensity values in (a) and (b). Storm durations are notably longer in California than in the Interior United States, which is a function of local climatology (see Staley et al., 2020). (d & e) Direct comparison between gauge and IMERG Peak 30min intensities. Values are weakly correlated.

distributions of DF and Non-DF intensity groups are statistically different as reported by the Mann-Whitney U test ($p \approx 0$). Despite strong class difference, IMERG 30min peak intensity values are underestimated in both regions, with the highest errors associated with storms in the Interior United

States (IUS), where short duration, high intensity convective events are observed to be a frequent mechanism for PFDF initiation (Cannon et al., 2008; Staley et al., 2020). Errors are lower in California (CA) where PFDFs are typically—although not exclusively—induced from short bursts of intense rainfall within longer duration atmospheric rivers (Kean et al., 2011; Oakley et al., 2017; Staley et al., 2020). This includes PFDFs resulting from convection-driven high intensity rainfall within a larger atmospheric river event (e.g., Oakley et al., 2018).

For both regions, storm duration plays an important role for effective precipitation estimate. The IMERG algorithm uses a Kalman filter to combine the more accurate, yet periodic, passive microwave (PMW) measurements with near constant infrared measurements from geostationary satellites (GEO-IR) to provide globally available precipitation data at 0.1° resolution in 30min intervals (Huffman et al., 2019). Product error is largely the result of the frequency of PMW overpasses (~12-16 times daily for a given location in CONUS) as temporally propagated PMW measurements are reported to be less correlated than calibrated GEO-IR after ± 90 minutes (Huffman et al., 2019; Joyce & Xie, 2011; Khan & Maggioni, 2019). As such, short duration storms may occur between PMW overpasses, leading to increased error. In contrast, longer storms allow for increased PMW overpasses for improved interpolation between individual measurements.

For further comparison, recurrence intervals (RIs) are frequently used to normalize variable intensity duration thresholds across different climatic regions. Staley et al. (2020) provide a detailed analysis of the RIs of PFDF generating storms across a variety of climatic regimes. However, a direct comparison of their RI data with IMERG-derived intensity values is challenging due to the methodology and available gauge records used in the NOAA Atlas 14 products (e.g., Perica et al., 2013) referenced by Staley et al (2020), and the comparatively short ~20-year record of IMERG data available for manual RI calculation. Nonetheless, a comparison of RI data for both sources yields

Standard Cross-Val		Staley et al. (2016) Data	*This Study				
Features			**# Model Runs**	**Mean AUC**	**Mean Threat Score**	**Train Size**	**Test Size**
dNBR/1000	PropHM23	Peak I15 mm/hr	50	0.81	0.43	499	125
dNBR/1000	PropHM23	Peak I30 mm/hr	50	0.8	0.45	495	124
dNBR/1000	PropHM23	IMERG Max 30min Intensity*	50	0.77	0.40	498	125
dNBR/1000*	PropHM23*	IMERG Max 30min Intensity*	50	0.74	0.37	494	124
Split: Train CA / Test IUS							
Features				**AUC**	**Threat Score**	**Train Size**	**Test Size**
dNBR/1000	PropHM23	Peak I15 mm/hr		0.7	0.37	467	154
dNBR/1000	PropHM23	Peak I30 mm/hr		0.68	0.35	465	157
dNBR/1000	PropHM23	IMERG Max 30min Intensity*		0.65	0.32	469	154
dNBR/1000*	PropHM23*	IMERG Max 30min Intensity*		0.58	0.27	469	123

Table 1. Results from the cross-validation process. This includes training on subsets of data collected from both regions (CA and IUS), and testing on a group not included in the training process. This process repeats for 50 unique train/test split iterations that preserve the original class balance of the dataset. Results using the train/test split from Staley et al. (2016) are included for a comparison to those reported by the previous study. We include the variable names used in their database for easier reference. “dNBR/1000” refers to the catchment averaged Differenced (or Delta) Normalized Burn Ratio. “PropHM23” is the proportion of slope pixels >23° burned at moderate to high severity to all slope pixels >23° within the catchment.

values of <5yrs, which suggests agreement with the general frequency of these storms (Staley et al. 2020) (See supplement).

We report model performance metrics in Table 1 and highlight the effect of IMERG 30min peak intensity values on model performance when used as the exclusive source of rainfall intensity data in the original USGS database. This provides the closest comparison for changes in model performance and potential utility as a data source when it stands in for gauge-collected threshold information. With both standard cross-validation and recreating the CA training/IUS testing split from Staley et al. (2016), model performance decreases from threat score values of 0.43-0.45 to 0.4, and 0.35-0.37 to 0.32, respectively. This suggests that while peak 30min intensity values are underestimated by the IMERG data, the relative differences between DF and non-DF classes are significant enough to build a model of comparable performance. Model results following this same validation process using all globally available satellite derived inputs are provided in the last row of each section in Table 1. Results are accordingly lower with the inclusion of burn severity and topographic metrics following our modified database recreation workflow. Differences here are largely the result of methodology: data provided by Staley et al. (2016) come from field-validated burn severity class data provided by the Burned Area Emergency Response (BAER) program. In contrast, we modify the methodology provided by Parks et al (2018) to rapidly approximate soil burn severity conditions using dNBR values provided by Landsat imagery. It is thus reasonable to assume that field validated burn severity data lead to a stronger model and Table 1 highlights the tradeoff in model skill between on-the-ground assessments and a fully automated method at the global scale.

While these results document our model performance within the scope of the original database, it is important to understand how the proposed framework would perform in the context of a near real time system, and especially on observations outside of the original training database within the United States. To demonstrate its applicability, we highlight model performance on data collected following the 2009 Beechworth Fire in Victoria, Australia, where Nyman et al. (2011, 2019) reported evidence of PFDF activity following heavy rainfall. The exact timing of the storms that initiated the majority of debris flows remain unreported; however, local news ties at least one PFDF inducing storm event to March 14th, 2009 which also led to a road closure (Dulhunty, 2009). As such, we map the location of a proximal debris flow following the wildfire recorded by Nyman et al. (2011) and reference it with the timing and nearby location of the events documented by Dulhunty (2009). We then apply the same modified data collection process used for the catchments referenced in the USGS database to gather the relevant catchment/storm characteristics for this event. The final model is trained on the fully recreated USGS database to make equivalent assessments for this area.

Figure 3 highlights PFDF probability as a function of daily IMERG max 30min rainfall intensity estimates from the approximate time of fire conclusion to several days past the time of recorded PFDF activity. As burn severity and slope properties are held constant during this time, changes in probability are dependent on rainfall. Notably the largest spike in model probability occurs during the same day which PFDF activity is recorded. Model probabilities remain consistently low for the rest of the days included in this time series. The threshold line at 31% in Figure 3 represents the mean probability for DF classification which maximized the Threat Score over 50 cross-validated model runs. In the event that the recorded debris flow occurred during this time, a system using this threshold would have correctly identified the PFDF in this area while avoiding false positives or false negatives during the prior month.

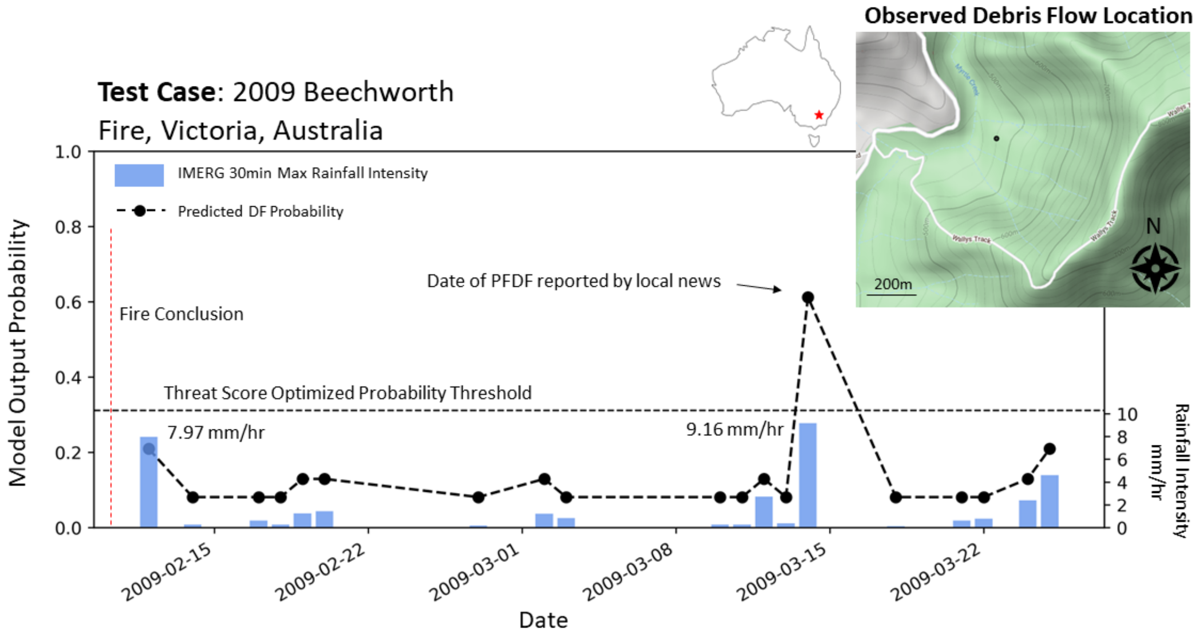

Figure 3. Daily PFDF probability as a function of IMERG-recorded 30min peak intensity rainfall for that day. Our model demonstrates nonlinear sensitivity to rainfall with a large jump in probability above 8.22 mm/hr , which is the median DF-inducing rainfall intensity value from our training dataset .The magnitude of this spike is nonetheless dependent on the burn and slope characteristics of the catchment (i.e., lower slope and burn severity values lead to probability increases at lower rates). Gauge-recorded accumulation values for this event vary between 12-40mm total (Dulhunty 2009). As the exact timing of this event was inferred from reporting by Nyman et al (2011, 2019) and Dulhunty (2009), there remains some uncertainty as to whether the debris flow occurred during the recorded storm time. Inset basemap image provided by Google Earth Engine.

4. Discussion

This work documents the development of a probabilistic PFDF initiation model that uses publicly available satellite-derived inputs. It is designed to operate in near real time and at the global scale. A significant difference between the proposed framework and that of others (e.g. Staley et al. 2016) is the intended use. Previous studies have focused on providing field-validated measurements of burn severity and gauge-collected rainfall in the Western United States to provide detailed insight into the physical constraints of PFDF initiation across spatially heterogenous terrain. Comparatively, our proposed framework employs an analogous remote sensing-based methodology that can provide routine situational awareness of these hazards. It does so by assessing the relevant combinations of slope, burn severity, and rainfall that may lead to the heightened likelihood of PFDF activity based on empirically correlated relationships of these characteristics derived from the Western United States. It appropriately demonstrates varied sensitivity to rainfall as a function of burn severity and slope parameters such that it produces reasonable model performance for its intended use cases.

As PFDFs are documented outside of the US (e.g, García-Ruiz et al., 2013; Jin et al., 2022; Nyman et al., 2019), model assessments to catchments on the global scale will represent an exciting opportunity for future research. Additional PFDF inventories would provide further insight as to which catchment characteristic thresholds are most relevant for PFDF activity in burned areas worldwide, with a strong focus on understanding which catchments simply do not have the appropriate characteristics to facilitate PFDF initiation. This is specifically an issue highlighted by Parise and Cannon (2008, 2012), who speculate that the lack of significant activity in other Mediterranean ecosystems might be a result of a combination of less severe fires, differences in watershed morphology, and/or patterns in rainfall. The proposed framework provides an opportunity to quantitively evaluate differences between these regions and those where PFDFs are more commonly observed. Similarly, each false alarm or miss provides valuable information that represents an opportunity to improve future hazard estimates. We note that the validation of future work may be limited to observations near populated areas—as debris flows occurring in remote regions often go unrecorded—but hope that this system will bring more attention to the locations and frequencies of these events where they are observed.

Despite adequate model performance, triggering rainfall intensity estimates in the IMERG dataset are indeed lower than their gauge-recorded counterpart. This behavior is expected. For instance, Khan et al. (2016) record lower IMERG performance in continental climates with dry summers compared to those with year-round precipitation. Further, performance of IMERG data has been extensively studied in several locations worldwide, where IMERG typically underestimates high intensity rainfall in complex terrain (e.g., Lu et al., 2019; Maggioni et al., 2017; Mayor et al., 2017; Nascimento et al., 2021; Navarro et al., 2020; Nepal et al., 2021; Rojas et al., 2021). Decreased IMERG performance due to orography is a known issue and remains an area for future work (Tan et al., 2019). Despite these limitations, recent versions of IMERG still help characterize patterns of orographic rainfall (e.g., Rojas et al., 2021; Sharma et al., 2020) and precipitation patterns on daily to seasonal timescales (e.g., Mayor et al., 2017; Nascimento et al., 2021; Nepal et al., 2021). Figure 2 demonstrates that PFDF triggering intensities are indeed higher than the peak intensities for the storms during which no PFDFs are recorded. This suggests that while the absolute magnitude of DF-inducing storm intensities will not be accurately captured, IMERG can differentiate these common—yet still significant—storms from the events during which PFDF initiation is statistically less likely.

5. Conclusion

We introduce a model for PFDF hazard assessment capable of near real time observations at the global scale. This model is trained on a modified database of historic PFDF activity and relies on globally available data products to increase situational awareness of these hazards globally. While the model does not increase lead-time for hazard warning, it is a resource for situational awareness where gauges and/or radar are not available. As our retrospective analysis demonstrates, this model can differentiate between DF and Non-DF inducing peak rainfall intensities across variable topographic and climatic regimes—even when hazard estimates are limited by the underestimations of intensity from IMERG data. There are still components we wish to incorporate into future iterations of the model, which include an updated training database with more events outside of the United States, and more recent data products such as IMERG Version 7 and Landsat 9. Once implemented, this model will provide a framework for near real time hazard assessments of PFDF processes on a global scale.

Acknowledgements

This research was supported by NASA's Disasters program through the solicitation for Earth Science Applications: Disaster Risk Reduction and Response (NNH18ZDA001N). We would like to thank Dennis Staley, Jason Kean, Matt Thomas, Francis Rengers and three anonymous reviewers for their valuable feedback on this work. We gratefully acknowledge the data provided by the USGS, which served as the basis for our model training, cited as Staley et al. (2016). Basemap data for Figures 1b and 3 provided by Stamen Design, under CC BY 3.0 (Data by OpenStreetMap, under ODbL) and Google Earth, respectively.

Data Availability Statement:

A developmental branch of the operational component of this work is available at the following Github repository: https://github.com/nasa/LHASA/tree/master/pfdf. Training data provided by the USGS is available at the following link: https://doi.org/10.3133/ofr20161106. Model training files are available at https://doi.org/10.5281/zenodo.7058363.

References

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1607171113

Bartos, M. (2020). *Pysheds: simple and fast watershed delineation in python.*

https://doi.org/10.5281/zenodo.3822494

Cannon, S. H., & DeGraff, J. (2009). The increasing wildfire and post-fire debris-flow threat in western USA, and implications for consequences of climate change. *Landslides - Disaster Risk Reduction*. https://doi.org/10.1007/978-3-540-69970-5_9

Cannon, S. H., Gartner, J. E., Wilson, R. C., Bowers, J. C., & Laber, J. L. (2008). Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California. *Geomorphology*, *96*(3–4), 250–269. https://doi.org/10.1016/j.geomorph.2007.03.019

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2939672.2939785

Debano, L. F. (2000). The role of fire and soil heating on water repellency in wildland environments: A review. *Journal of Hydrology*. https://doi.org/10.1016/S0022-1694(00)00194-3

Dennison, P. E., Brewer, S. C., Arnold, J. D., & Moritz, M. A. (2014). Large wildfire trends in the western United States, 1984-2011. *Geophysical Research Letters*. https://doi.org/10.1002/2014GL059576

Dulhunty, K. (2009, March 15). Road left covered in sludge. *The Border Mail*. https://www.bordermail.com.au/story/43678/road-left-covered-in-sludge/

Florsheim, J. L., Keller, E. A., & Best, D. W. (1991). Fluvial sediment transport in response to moderate storm flows following chaparral wildfire, Ventura County, southern California. *Geological Society of America Bulletin*. https://doi.org/10.1130/0016-7606(1991)103<0504:FSTIRT>2.3.CO;2

Florsheim, Joan L., Chin, A., O'Hirok, L. S., & Storesund, R. (2016). Short-term post-wildfire dry-ravel processes in a chaparral fluvial system. *Geomorphology*. https://doi.org/10.1016/j.geomorph.2015.03.035

Gabet, E. J. (2003). Sediment transport by dry ravel. *Journal of Geophysical Research: Solid Earth*. https://doi.org/10.1029/2001jb001686

García-Ruiz, J. M., Arnáez, J., Gómez-Villar, A., Ortigosa, L., & Lana-Renault, N. (2013). Fire-related debris flows in the Iberian Range, Spain. *Geomorphology*. https://doi.org/10.1016/j.geomorph.2012.03.032

Gartner, J. E., Cannon, S. H., & Santi, P. M. (2014). Empirical models for predicting volumes of sediment deposited by debris flows and sediment-laden floods in the transverse ranges of southern California. *Engineering Geology*, *176*, 45–56. https://doi.org/10.1016/j.enggeo.2014.04.008

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2017.06.031

He, H., & Ma, Y. (2013). Imbalanced learning: Foundations, algorithms, and applications. In *Imbalanced Learning: Foundations, Algorithms, and Applications*. https://doi.org/10.1002/9781118646106

Huffman, G., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., & Xie, P. (2019). NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) Prepared for: Global Precipitation Measurement (GPM) National Aeronautics and Space Administration (NASA). In *Algorithm Theoretical Basis*

Document (ATBD) Version 06. https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf

Jin, T., Hu, X., Liu, B., Xi, C., He, K., Cao, X., Luo, G., Han, M., Ma, G., Yang, Y., & Wang, Y. (2022). Susceptibility Prediction of Post-Fire Debris Flows in Xichang, China, Using a Logistic Regression Model from a Spatiotemporal Perspective. *Remote Sensing*, *14*(6), 1306. https://doi.org/10.3390/rs14061306

Joyce, R. J., & Xie, P. (2011). Kalman filter-based CMORPH. *Journal of Hydrometeorology*, *12*(6), 1547–1563. https://doi.org/10.1175/JHM-D-11-022.1

Kean, J. W., Staley, D. M., & Cannon, S. H. (2011). In situ measurements of post-fire debris flows in southern California: Comparisons of the timing and magnitude of 24 debris-flow events with rainfall and soil moisture conditions. *Journal of Geophysical Research: Earth Surface*. https://doi.org/10.1029/2011JF002005

Keeley, J. E. (2009). Fire intensity, fire severity and burn severity: A brief review and suggested usage. *International Journal of Wildland Fire*. https://doi.org/10.1071/WF07049

Key, C. H., & Benson, N. C. (2006). Landscape Assessment (LA) sampling and analysis methods. In *USDA Forest Service - General Technical Report RMRS-GTR*.

Khan, S., & Maggioni, V. (2019). Assessment of Level-3 Gridded Global Precipitation Mission (GPM) Products Over Oceans. *Remote Sensing*, *11*(3), 255. https://doi.org/10.3390/rs11030255

Khan, S., Maggioni, V., & Porcacchia, L. (2016). Uncertainties associated with the IMERG Multi-Satellite precipitation product. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2016-Novem*, 2127–2130. https://doi.org/10.1109/IGARSS.2016.7729549

Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-Triggered Landslide

Hazard for Situational Awareness. *Earth's Future*, *6*(3), 505–523. https://doi.org/10.1002/2017EF000715

Lamb, M. P., Levina, M., Dibiase, R. A., & Fuller, B. M. (2013). Sediment storage by vegetation in steep bedrock landscapes: Theory, experiments, and implications for postfire sediment yield. *Journal of Geophysical Research: Earth Surface*. https://doi.org/10.1002/jgrf.20058

Lamb, M. P., Scheingross, J. S., Amidon, W. H., Swanson, E., & Limaye, A. (2011). A model for fire-induced sediment yield by dry ravel in steep landscapes. *Journal of Geophysical Research: Earth Surface*. https://doi.org/10.1029/2010JF001878

Letey, J. (2001). Causes and consequences of fire-induced soil water repellency. *Hydrological Processes*. https://doi.org/10.1002/hyp.378

Liu, T., McGuire, L. A., Oakley, N., & Cannon, F. (2022). Temporal changes in rainfall intensity–duration thresholds for post-wildfire flash floods in southern California. *Natural Hazards and Earth System Sciences*, *22*(2), 361–376. https://doi.org/10.5194/nhess-22-361-2022

Lu, X., Tang, G., Wang, X., Liu, Y., Jia, L., Xie, G., Li, S., & Zhang, Y. (2019). Correcting GPM IMERG precipitation data over the Tianshan Mountains in China. *Journal of Hydrology*, *575*(January), 1239–1252. https://doi.org/10.1016/j.jhydrol.2019.06.019

Maggioni, V., Nikolopoulos, E. I., Anagnostou, E. N., & Borga, M. (2017). Modeling satellite precipitation errors over mountainous terrain: The influence of gauge density, seasonality, and temporal resolution. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2017.2688998

Mayor, Y. G., Tereshchenko, I., Fonseca-Hernández, M., Pantoja, D. A., & Montes, J. M. (2017). Evaluation of error in IMERG precipitation estimates under different topographic conditions and temporal scales over Mexico. *Remote Sensing*, *9*(5), 1–18. https://doi.org/10.3390/rs9050503

Miller, J. D., & Safford, H. (2012). Trends in wildfire severity: 1984 to 2010 in the Sierra Nevada, Modoc Plateau, and southern Cascades, California, USA. *Fire Ecology*. https://doi.org/10.4996/fireecology.0803041

Mueller, S. E., Thode, A. E., Margolis, E. Q., Yocom, L. L., Young, J. D., & Iniguez, J. M. (2020). Climate relationships with increasing wildfire in the southwestern US from 1984 to 2015. *Forest Ecology and Management*. https://doi.org/10.1016/j.foreco.2019.117861

NASA JPL. (2020). *NASA DEM*. NASA EOSDIS Land Processes DAAC. https://doi.org/https://doi.org/10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001

Nascimento, J. G., Althoff, D., Bazame, H. C., Neale, C. M. U., Duarte, S. N., Ruhoff, A. L., & Gonçalves, I. Z. (2021). Evaluating the latest imerg products in a subtropical climate: The case of paraná state, brazil. *Remote Sensing*, *13*(5), 1–20. https://doi.org/10.3390/rs13050906

Navarro, A., García-Ortega, E., Merino, A., Sánchez, J. L., & Tapiador, F. J. (2020). Orographic biases in IMERG precipitation estimates in the Ebro River basin (Spain): The effects of rain gauge density and altitude. *Atmospheric Research*, *244*(February), 105068. https://doi.org/10.1016/j.atmosres.2020.105068

Nepal, B., Shrestha, D., Sharma, S., Shrestha, M. S., Aryal, D., & Shrestha, N. (2021). Assessment of GPM-Era satellite products' (IMERG and GSMaP) ability to detect precipitation extremes over mountainous country nepal. *Atmosphere*, *12*(2). https://doi.org/10.3390/atmos12020254

Nikolopoulos, E. I., Destro, E., Bhuiyan, M. A. E., Borga, M., & Anagnostou, E. N. (2018). Evaluation of predictive models for post-fire debris flow occurrence in the western United

States. *Natural Hazards and Earth System Sciences*, *18*(9), 2331–2343. https://doi.org/10.5194/nhess-18-2331-2018

Nyman, P., Rutherfurd, I. D., Lane, P. N. J., & Sheridan, G. J. (2019). Debris flows in southeast Australia linked to drought, wildfire, and the El Niño-Southern Oscillation. *Geology*, *47*(5), 491–494. https://doi.org/10.1130/G45939.1

Nyman, P., Sheridan, G. J., Smith, H. G., & Lane, P. N. J. (2011). Evidence of debris flow occurrence after wildfire in upland catchments of south-east Australia. *Geomorphology*, *125*(3), 383–401. https://doi.org/10.1016/j.geomorph.2010.10.016

Nyman, P., Smith, H. G., Sherwin, C. B., Langhans, C., Lane, P. N. J., & Sheridan, G. J. (2015). Predicting sediment delivery from debris flows after wildfire. *Geomorphology*, *250*, 173–186. https://doi.org/10.1016/j.geomorph.2015.08.023

Oakley, N. S., Cannon, F., Munroe, R., Lancaster, J. T., Gomberg, D., & Martin Ralph, F. (2018). Brief communication: Meteorological and climatological conditions associated with the 9 January 2018 post-fire debris flows in Montecito and Carpinteria, California, USA. *Natural Hazards and Earth System Sciences*. https://doi.org/10.5194/nhess-18-3037-2018

Oakley, N. S., Lancaster, J. T., Kaplan, M. L., & Ralph, F. M. (2017). Synoptic conditions associated with cool season post-fire debris flows in the Transverse Ranges of southern California. *Natural Hazards*. https://doi.org/10.1007/s11069-017-2867-6

Parise, M., & Cannon, S. H. (2008). *The effects of wildfires on erosion and debris-flow generation in Mediterranean climatic areas: a first database. Proceedings of 1st World Landslide Forum*.

Parise, M., & Cannon, S. H. (2012). Wildfire impacts on the processes that generate debris flows in burned watersheds. *Natural Hazards*. https://doi.org/10.1007/s11069-011-9769-9

Parks, S. A., Holsinger, L. M., Voss, M. A., Loehman, R. A., & Robinson, N. P. (2018). Mean composite fire severity metrics computed with google earth engine offer improved accuracy and expanded mapping potential. *Remote Sensing*, *10*(6), 1–15. https://doi.org/10.3390/rs10060879

Perica, S., Martin, D., Pavlovic, S., Roy, I., Laurent, M. St., Trypaluk, C., Unruh, D., Yekta, M., & Bonnin, G. (2013). *Precipitation-Frequency Atlas of the United States Volume 8 Version 2.0: Midwestern States*. https://www.weather.gov/media/owp/oh/hdsc/docs/Atlas14_Volume8.pdf

Rojas, Y., Minder, J. R., Campbell, L. S., Massmann, A., & Garreaud, R. (2021). Assessment of GPM IMERG satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central Chile. *Atmospheric Research*, *253*(December 2020), 105454. https://doi.org/10.1016/j.atmosres.2021.105454

Sharma, S., Chen, Y., Zhou, X., Yang, K., Li, X., Niu, X., Hu, X., & Khadka, N. (2020). Evaluation of GPM-Era satellite precipitation products on the southern slopes of the central Himalayas against rain gauge data. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111836

Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., & Laber, J. L. (2013). Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides*, *10*(5), 547–562. https://doi.org/10.1007/s10346-012-0341-9

Staley, D. M., Kean, J. W., & Rengers, F. K. (2020). The recurrence interval of post-fire debris-flow generating rainfall in the southwestern United States. *Geomorphology*, *370*, 107392. https://doi.org/10.1016/j.geomorph.2020.107392

Staley, D. M., Negri, J. A., Kean, J. W., Laber, J. L., Tillery, A. C., & Youberg, A. M. (2016). *Updated logistic regression equations for the calculation of post-fire debris-flow likelihood in the western United States*.

2016–1106. https://doi.org/10.3133/ofr20161106

Staley, D. M., Negri, J. A., Kean, J. W., Laber, J. L., Tillery, A. C., & Youberg, A. M. (2017). Prediction of spatially explicit rainfall intensity–duration thresholds for post-fire debris-flow generation in the western United States. *Geomorphology*. https://doi.org/10.1016/j.geomorph.2016.10.019

Tan, J., Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2019). IMERG V06: Changes to the morphing algorithm. *Journal of Atmospheric and Oceanic Technology*, *36*(12), 2471–2482. https://doi.org/10.1175/JTECH-D-19-0114.1

Tang, H., McGuire, L. A., Rengers, F. K., Kean, J. W., Staley, D. M., & Smith, J. B. (2019). Developing and Testing Physically Based Triggering Thresholds for Runoff-Generated Debris Flows. *Geophysical Research Letters*, *46*(15), 8830–8839. https://doi.org/10.1029/2019GL083623

U.S. Geological Survey. (n.d.). *Landsat 7 Collection 1 Tier 1 Surface Reflectance*. U.S. Geological Survey. https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-1

Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and earlier spring increase Western U.S. forest wildfire activity. *Science*. https://doi.org/10.1126/science.1128834